

## PATENT ABSTRACTS OF JAPAN

(11)Publication number : 10-326154

(43)Date of publication of application : 08.12.1998

(51)Int.Cl.

G06F 3/06

G06F 3/06

(21)Application number : 09-133903

(71)Applicant : BROTHER IND LTD

(22)Date of filing : 23.05.1997

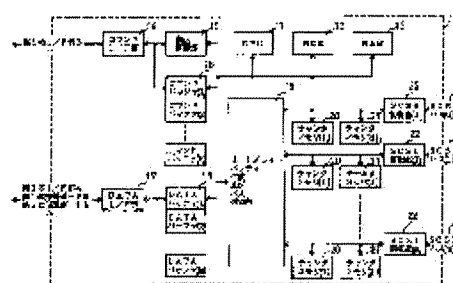
(72)Inventor : YASUDA MASARU

## (54) ARRAY CONTROLLER AND DATA SERVER

## (57)Abstract:

**PROBLEM TO BE SOLVED:** To provide an array controller that can fast transfer data to a host computer.

**SOLUTION:** This controller has an I/F part 14, which transfers at least a control command or status to a central control part and a DATA I/F part 17 that is prepared separately from the part 14 and transfers the data, which have been read out of a hard disk 7. Thus, the I/F part 14 which transfers the relative information on the control command, etc., is prepared separately from the I/F part 17 which reads the data out of the disk 7 to transfer them. Therefore, the controller can receive another control command from the central control part even when it's transferring the data to an external device, for example. In other words, the transfer of data of the I/F part 14 is never affected even if the I/F part 17 is kept in an occupied state from the start of the transfer of data through the end of the transfer of data.



**\* NOTICES \***

JPO and INPIT are not responsible for any damages caused by the use of this translation.

- 1.This document has been translated by computer. So the translation may not reflect the original precisely.
- 2.\*\*\*\* shows the word which can not be translated.
- 3.In the drawings, any words are not translated.

---

**CLAIMS**

---

[Claim(s)]

[Claim 1]An array controller which can transmit data which is controllable in parallel and read two or more sets of memory storage characterized by comprising the following from said memory storage based on directions from an external host computer to an external device containing said host computer.

An interface part for commands for transmitting control-commands pertinent information between said host computers.

An interface part for data for performing a data transfer which it was provided separately from the interface part for commands concerned, and was read from said memory storage to said external device.

[Claim 2]The array controller according to claim 1 constituting based on directions from said host computer so that writing of data to said memory storage may also be possible and said interface part for data may be used also for a data transfer for writing from said external device.

[Claim 3]A data server comprising:

The array controller according to claim 1 or 2.

Have the 2nd interface part for connecting with two or more memory storage controlled by this array controller, the 1st interface part for connecting with said interface part for commands of said array controller, and said interface part for data, and. A host computer it can be directed that data according to a request from a terminal unit of the data supply point which is said external device is read, and transmits to said terminal unit to an array controller of said storage device array.

[Claim 4]A data server comprising:

The array controller according to claim 2.

Have the 2nd interface part for connecting with two or more memory storage controlled by this array controller, the 1st interface part for connecting with said interface part for commands of said array controller, and said interface part for data, and. Can direct to read data according to a request from a terminal unit of the data supply point which is said external device, and to transmit to said terminal unit to an array controller of said storage device array, and. A host computer it can be directed that writes data for writing transmitted from said external device in said memory storage.

[Claim 5]As said 1st interface part by the side of a host computer for connecting with an interface part for commands of said array controller, The data server according to claim 3 or 4 using an interface part or an RS232C interface part for printers with which a host computer is provided.

---

[Translation done.]

**\* NOTICES \***

JP0 and INPIT are not responsible for any damages caused by the use of this translation.

- 1.This document has been translated by computer. So the translation may not reflect the original precisely.
- 2.\*\*\*\* shows the word which can not be translated.
- 3.In the drawings, any words are not translated.

---

**DETAILED DESCRIPTION**

---

[Detailed Description of the Invention]

[0001]

[Field of the Invention]This invention is operating two or more memory storage in parallel, and a vast quantity of data of an image, a sound, etc. processed, for example with computer apparatus is made to record on a high speed and real time, It is related with the data server provided with the array controller which can be made to read, and its array controller.

[0002]

[Description of the Prior Art]From before memory storage, such as a hard disk drive, in operating two or more set parallel, for example. The storage device array which can respond to a high speed and real time is known by the demand of the data read from a main computer, or data writing, and the array controller used as the control section is also known. The data server using this storage device array is also considered.

[0003]And to the conventional array controller by the personal computer (it is hereafter written as a personal computer.) as a host computer, or the interface part for workstations. There is a thing of the gestalt of the add-in board provided with the interface part called PCI (Peripheral Component Interconnect). The array controller of this add-in board gestalt is inserted in the expansion slot by the side of a personal computer or a workstation, the array controller itself is connecting with the memory storage group of a controlled object, and it performs control to a memory storage group based on the directions from a personal computer etc.

[0004]

[Problem(s) to be Solved by the Invention]By the way, in the interface part between the personal computer as a host device, a workstation, etc. and an array controller, it is necessary to transmit the data for read-out memorized by the memory storage group, and the data for writing for recording on a memory storage group at high speed. It is necessary to transmit various control commands for making the reading processing and writing processing of data perform besides the read data or the data for writing, the status to which it is answered from a processing object as a result which receives control commands, etc.

[0005]However, generally the data for read-out memorized by the memory storage group has large data size so that it may be represented by the video data, for example, and predetermined time is needed for the data transfer. Therefore, during the data transfer, directions of data read with a new host computer cannot be issued at an array controller. That is, it will wait until data communications are completed, and a response falls. And this response falls as the frequency where a host computer issues data read directions becomes high.

[0006]Therefore, when the request from many terminals tends to be received when it applies to a data server, and it is going to supply the data according to the request to real time as much as possible. The fall of the response between the above-mentioned host computer and an array controller leads to the fall of the response of data supply to the request from a terminal. For example, the video-on-demand (VOD) service which enabled it to see the movie etc. which are wished to have when a user wishes is studied, and this VOD service, If desired video televising is requested from the set top box (STB) provided in the user side who becomes an information output side, the image and sound according to the request will be transmitted to a user from the

information feed unit called the video server prepared for the purveyor-of-service side. In such a system, the information for supply outputted to the information printing terminal side is video information, and is outputted at any time according to the request from an information printing terminal. Therefore, if the control commands for the next processing cannot be published, either, unless the data transfer processing to one certain request is completed, a response will fall, and the fall of this response will lead to the deterioration of information supply service.

[0007] This invention is made in order to solve the problem mentioned above, and it sets it as the 1st purpose to provide the array controller in which the high-speed data transfer between host computers is possible. When it applies to a video server etc., offer of an effective data server is set as the 2nd purpose in respect of the improvement in response time.

[0008]

[The means for solving a technical problem and an effect of the invention] The invention according to claim 1 made in order to attain said 1st purpose, In the array controller which can transmit the data which is controllable in parallel and read two or more sets of memory storage from memory storage based on the directions from an external host computer to the external device containing a host computer, The interface part for commands for transmitting control-commands pertinent information at least between host computers, It is provided separately from the interface part for commands, and has an interface part for data for performing the data transfer read from memory storage to the external device. Information like the status to which it is answered from the controlled object side other than the control commands themselves as a result of the control commands is included in control-commands pertinent information.

[0009] Although the array controller of this invention can transmit the data read from memory storage to the external device containing a host computer based on the directions from an external host computer, About the exchange of the control-commands pertinent information between host computers, for example, control commands, and status, it carries out via the interface part for commands, and performs transmitting the data read from memory storage to an external device via the interface part for data.

[0010] Thus, by preparing individually an interface for transmission of control-commands pertinent information, such as control commands and status, and an interface for data transfer in a case of reading data from memory storage, respectively, For example, even if it is during data transfer to an external device, another control commands from a host computer can be received, and an array controller which can realize data transfer high-speed as a whole can be obtained. That is, since an interface for control-commands transmission and an interface for data transfer have dissociated, Even if an occupied state of an interface part for data for data transfer from a data transfer start to an end arises, An interface part for commands is not influenced, but it can perform a control-commands control section continuously between host computers, without interrupting control management based on a predetermined program.

[0011] Not only data read and transmission from memory storage but also constituting writing of data to memory storage possible based on directions from a host computer, and constituting an interface part for data so that it may use also for a data transfer for writing from an external device is considered. If it does in this way, the data itself which is read from memory storage and transmitted to an external device can be updated and added, and high-speed data transfer can be realized at the time of the writing.

[0012] An invention shown in claim 3 made on the other hand in order to attain the 2nd purpose, An array controller mentioned above and two or more memory storage controlled by an array controller, Have the 2nd interface part for connecting with the 1st interface part for connecting with an interface part for commands of an array controller, and an interface part for data, and. It is a data server provided with a host computer it can be directed that data according to a request from a terminal unit of the data supply point which is an external device is read, and transmits to a terminal unit to an array controller of a storage device array.

[0013] Although this data server is applicable to what is called a video server etc., it leads to improvement of response time to a supply demand of a video data carried out from a terminal unit of the data supply point in that case. In this case, although it is possible that a request from two or more terminal units is advanced by arbitrary timing, If a host computer cannot take out

directions to an array controller to the 2nd request unless data transfer according to the first request is completed, the data transfer according to the 2nd request itself will be overdue. Similarly, a response as the whole -- 3rd henceforth will be overdue -- will fall. Therefore, as mentioned above, by keeping data transfer from an exchange between an array controller and a host computer being influenced, a response in the portion improves and a response to a data supply demand carried out from an external terminal unit also improves.

[0014]When it applies to a video server especially, it is possible to be used for a system of video-on-demand (VOD) service which enabled it to see a movie etc. which are wished to have when a user wishes. In such a system, since it is what outputs data at any time according to a request from a user's terminal unit, it leads to the progression in quality of data supply service that improvement in a response to a data supply demand can be aimed at.

[0015]What is necessary is just to constitute so that directions with writing data for writing to which a host computer is transmitted from an external device in memory storage are possible in using an array controller which can also perform data writing control to memory storage, as shown in claim 4. Since an exchange between an array controller and a host computer is not similarly influenced by data transfer in this case, even if a response in that portion improves and writing operation of data for writing to memory storage, etc. occur, a response to a data supply demand carried out from an external terminal unit after all improves.

[0016]When realizing as such a data server, as the 2nd interface part for data transfer by the side of a host computer, I/F for PCI explained also as conventional technology can be used. It connects by introducing this into expansion slots, such as a personal computer, and is made data transfer. It is possible to use an interface part or an RS232C interface for printers with which a host computer is provided, for example as the 1st interface by the side of a host computer corresponding to an interface part for commands for transmission of control-commands pertinent information. Of course, although an interface part for exclusive use may be provided, since there are many computers which it has from the first if it is an interface part or an RS232C interface for printers, it can be used.

[0017]

[Embodiment of the Invention]Hereafter, an embodiment of the invention is described with reference to drawings. Drawing 1 is a block diagram showing the composition at the time of [ whole ] applying to a video server. Two or more SCSI buses 6 are connected to the array controller 5. In distinguishing two or more SCSI buses 6, it describes as a SCSI bus (0) - a SCSI bus (m). Two or more hard disk drives (HDD) are connected to each SCSI bus 6 of the. In distinguishing two or more hard disk drives 7, suppose that ID (0) - ID (n) are attached and described. The disk array A is constituted by two or more hard disk drives 7 connected to these array controllers 5, two or more SCSI buses 6, and it.

[0018]On the other hand, the host computer B which controls this disk array A is constituted by the storage parts store 1, the central controlling part 2, the 1st interface (it outlines the following "I/F") part 3, the 2nd I/F part 4, the 1st add-in board 8, the 2nd add-in board 9, etc.

[0019]The 1st I/F part 3 is an interface for performing the control commands as "control-commands pertinent information", and transmission of status, and the 2nd I/F part 4 is an interface for mainly performing data transfer. The array controller 5 mentioned above is connected to the central controlling part 2 via these two I/F parts 3 and 4.

[0020]The 1st add-in board 8 and 2nd add-in board 9 are connected to the 2nd I/F part 4. These two add-in boards 8 and 9 receive the data written in the hard disk drive 7 of the disk array A from the outside, and transmit the received data to the array controller 5, and. It is a LAN board for transmitting to the terminal unit (not shown) etc. which were connected via the communication line which does not illustrate the data recorded on the buffer which it is read from the hard disk drive 7, and is later mentioned in the array controller 5.

[0021]Although the central controlling part 2 of the host computer B serves as a control subject of the main operations as a video server, it roughly divides and there are the following two kinds. When a request is advanced from the terminal unit (not shown) connected via the communication line (not shown) accommodated in the 1st add-in board 8 or 2nd add-in board 9, one of them, It is operation of transmitting the data according to a request to the terminal unit,

and another is operation of incorporating a sound and picture image data from the exterior newly, and storing it in the hard disk drive 7 of the disk array A.

[0022]Data distribution to an external terminal unit is performed as follows. The central controlling part 2 as a control subject of a video server, As opposed to the supply request of the image and voice data which were sent from STB (Set Top Box) as a terminal unit, etc. for example, it was connected via the 1st add-in board 8 and 2nd add-in board 9, The data according to the request memorized as a file is read to the hard disk drive 7 via the array controller 5, and the terminal unit side is supplied via the 1st add-in board 8 and 2nd add-in board 9.

[0023]On the other hand, data storage operation is performed as follows to the hard disk drive 7. A new sound and picture image data are incorporated from the exterior via the 1st add-in board 8 and 2nd add-in board 9, or are incorporated via the I/F part etc. which were connected to the 2nd I/F part 4 and which are not illustrated, and it is transmitted to the array controller 5. The array controller 5 records the transmitted data on the hard disk drive 7 in the form of a file etc.

[0024]Next, the array controller 5 is explained with reference to drawing 2. The array controller 5 is provided with the following.

CPU11 which manages control of the array controller 5 whole.

ROM12 a program, various parameters, etc. were remembered to be.

RAM13 used as the work area of CPU11, and data storage area.

The command I/F part 14 equivalent to "the I/F part for commands", and the interruption control section 15, Two or more command buffers 16 and the DATAI/F part 17 equivalent to "the I/F part for data", Two or more DATA buffers 18, read/write parity generating and the bus changeover section 19, two kinds of chunk memories, i.e., chunk memory (1) 20, and chunk memory (2) 21, and the SCSI control section 22.

Chunk memory (1) 20 and chunk memory (2) 21, and the SCSI control section 22, Since it is provided every SCSI bus 6 and SCSI bus 6 exists to ID(0) - (m) in this embodiment, the SCSI control section 22 will exist to ID(0) - (m) similarly, and chunk memory (1) 20 and chunk memory (2) 21 will be connected to each.

[0025]It is connected to the central controlling part 2 via the 1st I/F part 3 (refer to drawing 1) by the side of the host computer B, and it is connected to the command buffer 16 and the interruption control section 15 in the array controller 5, and said command I/F part 14 is used for transmission of control-commands pertinent information. Information like the status to which it is answered from the controlled object side other than the control commands themselves as a result of those control commands is included in this "control-commands pertinent information."

[0026]The command buffer 16 is prepared by two or more area so that one or more commands can be processed simultaneously. In distinguishing two or more command buffers 16, suppose that the number of 1 - h is attached and described like a command buffer (1) - a command buffer (h).

[0027]The interruption control section 15 receives CPU11, when the central controlling part 2 generates interruption to the array controller 5, after setting a command to the command buffer 16, When the array controller 5 generates interruption to the central controlling part 2, it has the function to generate interruption to the central controlling part 2, respectively, via the 1st I/F part 3.

[0028]When applying what is called a personal computer as the host computer B, it is also possible to apply general-purpose ports, such as a printer port and an RS232C port, and to, apply I/F for exclusive use, of course, for example as the 1st I/F part 3.

[0029]On the other hand, the DATAI/F part 17 is connected to the central controlling part 2 via the 2nd I/F part 4 (refer to drawing 1) by the side of the host computer B, and it is connected to the 1st add-in board 8 and the 2nd add-in board 9 grade, and. It is connected to the DATA buffer 18 in the array controller 5, and is mainly used at the time of a data transfer.

[0030]The DATA buffer 18 is prepared by two or more area so that the demand from two or more add-in boards 8 and 9 can be met. In distinguishing two or more DATA buffers 18, suppose that the number of 1 - k is attached and described like a DATA buffer (1) - a command buffer (k).

[0031]As this DATA/F part 17, the art of PCI (Peripheral Component Interconnect) explained also as conventional technology can be used. That is, it is an add-in board gestalt, and array controller 5 the very thing will insert the array controller 5 of this add-in board gestalt in the expansion slot by the side of the host computer B, and will connect. Of course, the 2nd I/F part 4 by the side of the host computer B in this case is also I/F corresponding to PCI.

[0032]Next, operation of read/write parity generating and the bus changeover section 19 is explained. Read/write parity generating and the bus changeover section 19 are provided with internal bus side I/O control unit 30, parity generating, the check part 31, and SCSI bus side I/O control unit 32 as shown in drawing 3 – drawing 6.

[0033]An internal bus is assumed to be 32 bit width, it treats as a bus of the unit summarized every 8 bits, the number of 0–3 for discernment of each bus is attached, and it expresses with this embodiment the A bus 0, the A bus 1, the A bus 2, and the A bus 3. The SCSI bus connected to the array controller 5 is made into five 8-bit SCSI buses, and one line of them is explained as what is other four buses for parity.

[0034]The operation at the time of the light to the hard disk drive 7 is explained to the beginning. As shown in drawing 3, internal bus side I/O control unit 30 at this time serves as an input to the A buses 0–3, and serves as an output to B buses 0–3. In detail, the contents of the A bus 3 will be outputted to the A bus 2 and B bus 3 to B bus 0 as it is at the A bus 0 and B bus 1 at the A bus 1 and B bus 2. The value of four B buses 0–3 is inputted into parity generating and the check part 31, and the parity account value of the value which is four is outputted to B bus P. In SCSI bus side I/O control unit 32, the value of B bus P and B buses 0–3 is inputted, and it is outputted to C bus 0–4.

[0035]Next, the operation at the time of the usual lead of the hard disk drive 7 is explained with reference to drawing 4. Since SCSI bus side I/O control unit 32 is operation of the lead from the hard disk drive 7, the C bus 0–4 side serves as an input, and the value of the bus is outputted to B busP and B buses 0–3. Since the value of B busP and B buses 0–3 is inputted, by taking the exclusive OR of the value of B buses 0–3, a parity value is calculated in parity generating and the check part 31, and the result compared with the value of B bus P is outputted to them as a parity check result. B buses 0–3 are inputted into internal bus side I/O control unit 30, and it outputs to the A buses 0–3.

[0036]Next, the hard disk drive 7 connected to the bus of the identification number 3 breaks down, From the hard disk drive 7 connected to the bus of the identification number 0, and the hard disk drive 7 connected to other three buses (identification numbers 1–3), operation in case the data of the hard disk drive 7 connected to the bus of the identification number 4 is restored is explained with reference to drawing 5.

[0037]In this case, the value of C bus 0–3 is inputted into SCSI bus side I/O control unit 32, and the value of C bus 4 is unfixed. To the B bus 0–3 side, the value of C bus 0–3 is outputted to B busP and B buses 0–2, respectively. And the value of B busP and B buses 0–2 is inputted into parity generating and the check part 31, the value which should be outputted to C bus 4 by taking the exclusive OR of those values is restored, and it is outputted to B bus 3. B buses 0–3 are inputted into internal bus side I/O control unit 30 like the case of drawing 4, and it is outputted to the A buses 0–3 as it is.

[0038]Next, the hard disk drive 7 connected to the bus of the identification number 4 breaks down, After being exchanged for the hard disk drive 7 with the broken normal hard disk drive 7, with the data from the hard disk drive 7 connected to the bus of the identification numbers 0–3. Operation in case the data of the hard disk drive 7 connected to the bus of the identification number 4 is reconstructed (Ribi Rudo) is explained with reference to drawing 6.

[0039]The value of C bus 0–3 is inputted into SCSI bus side I/O control unit 32, and the value of C bus 0–3 is outputted to it to B busP and B buses 0–2 to the B bus 0–3 side, respectively. The value of B busP and B buses 0–2 is inputted into parity generating and the check part 31, the data memorized by the hard disk drive 7 connected to C bus 4 by taking the exclusive OR of those values is restored, and it is outputted to B bus 3.

[0040]In SCSI bus side I/O control unit 32, unlike other buses, only B bus 3 serves as an input, and data is outputted to C bus 4. It becomes possible to write the data reproduced by this in the

hard disk drive 7, and to reconstruct the data of the completely same contents as the broken hard disk drive 7. It also becomes possible to lead simultaneously, reconstructing the data memorized by the hard disk drive 7 by considering the B bus 0-3 side of internal bus side I/O control unit 30 as an input, and outputting that value from the A bus 0-3 side at this time.

[0041]Transmission of lead data and right data is performed via read/write parity generating and the bus changeover section 19 by the above-mentioned operation between the DATA buffer 18, the 1st chunk memory 20, or the 2nd chunk memory 21 (refer to [drawing 1](#)).

[0042]Next, data is explained from the chunk memories 20 and 21 to the hard disk drive 7 about transfer operation and its operation which transmits data to the chunk memories 20 and 21 from the hard disk drive 7 conversely. It mentioned above -- as -- each SCSI bus 6 -- the hard disk drive 7 -- 1 -- or it is connected by more than one. The SCSI control section 22 controls SCSI bus 6, and at the time of writing. Chunk memory (1) 20 and chunk memory (2) 21. The data written in 21 is transmitted to the hard disk drive 7, and the data read from the hard disk drive 7 is transmitted to chunk memory (1) 20 or chunk memory (2) 21 at the time of read-out. The reason with two chunk memories is for accelerating access to the hard disk drive 7.

[0043]Hereafter, the data flow at the time of the time of writing and read-out is explained collectively. At the time of writing, a write data is transmitted to one of the DATA buffers 18 among the prepared DATA buffers 18 via the DATA/F part 17. From the DATA buffer 18, a write data is transmitted to chunk memory (1) 20 via a read/write parity generating part and the bus changeover section 19. Data is transmitted and memorized by the hard disk drive 7 connected to SCSI bus 6 via the SCSI control section 22 from chunk memory (1) 20 after transmission is completed. Simultaneously with this data transfer, the data which should be transmitted to another chunk memory (2) 21 hard disk drive 7 next is beforehand transmitted from the DATA buffer 18. By this, after data transfer ending from one chunk memory to the hard disk drive 7, without opening useless time, the data of another chunk memory can be immediately transmitted to the hard disk drive 7, and improvement in the speed of write operation is attained.

[0044]At the time of read-out, data is transmitted to the reverse in order of hard disk drive 7 -> SCSI bus 6 -> SCSI control-section 22 -> chunk memory (1) 20 or chunk memory (2) 21. The following read data is transmitted from the hard disk drive 7 to the vacant chunk memories 20 and 21 during the data transfer from the chunk memories 20 and 21 to the DATA buffer 18 at the time of read-out as well as the time of writing. By it, read operation is accelerable.

[0045]Next, operation when there is a transfer request of a file from the two add-in boards 8 and 9 is explained with reference to the sequence diagram of [drawing 7](#) and [drawing 8](#). [Drawing 7](#) is a case of the array controller 5 of this embodiment, and is an example in the case of performing data transfer via command I/F14 and DATA/F17 which are shown in [drawing 2](#) between the 1st I/F part 3 by the side of the host computer B, and the 2nd I/F part 4, respectively. On the other hand, [drawing 8](#) is an example in case only one has I/F, therefore the array controller 5 performs data transfer in one I/F in between with the central controlling part 2 for comparison.

[0046]First, with reference to [drawing 8](#), the case of the comparative example to which the array controller 5 and the central controlling part 2 carry out data transfer in one I/F is explained. As shown in (1) among [drawing 8](#), transmission of the file A is first requested to the central controlling part 2 from the 1st add-in board 8. As shown in (2) among [drawing 8](#), the central controlling part 2 publishes the lead demand of the file A to the array controller 5.

[0047]Then, as shown in (3) among [drawing 8](#), within the array controller 5, According to the data read of the file A, and said operation procedures, data is transmitted to the DATA buffer 18 from the hard disk drive 7 via the SCSI control section 22, the chunk memories 20 and 21, read/write parity generating, and the bus changeover section 19 to the demand.

[0048]As shown in (4) among [drawing 8](#), the array controller 5 notifies the end of a lead of the file A to the central controlling part 2 after the end of a file lead. And by the central controlling part 2, as shown in (5) among [drawing 8](#), after receiving the end of a file A lead, the transfer request of the file A is published to the 1st add-in board 8. And as shown in (6) among [drawing 8](#), in the 1st add-in board 8, the transfer start of the file A is notified to the central controlling part 2 to the demand.



[0049]As shown in (7) immediately after this and among drawing 8, the request to the file B occurs from the 2nd add-in board 9, and it is notified to the central controlling part 2. However, since it is after the transfer start status of the file A was notified by the 1st add-in board 8 to the central controlling part 2 at this time, as shown in (8) among drawing 8, the 1st add-in board 8 starts transmission of the file A.

[0050]As for under the data transfer by the 1st add-in board 8 started by (8) among drawing 8, the bus is used for the data transfer from the DATA buffer 18 in the array controller 5 to the add-in board 8. For this reason, the central controlling part 2 cannot immediately publish the request of the file B from the 2nd add-in board 9. Therefore, as shown in (9) among drawing 8, the central controlling part 2 will notify the lead demand of the file B to the array controller 5 for the first time, after transmission of the file A from the array controller 5 to the 1st add-in board 8 is completed.

[0051]And as shown in (10) among drawing 8, in the central controlling part 2, the transmission termination status of the file A is received from the 1st add-in board 8. As shown in (11) among drawing 8, with the array controller 5, file read operation is started from the hard disk drive 7 to the lead demand of the file B, and, almost simultaneously with it, the data of the file B read into the DATA buffer 18 from the hard disk drive 7 is transmitted.

[0052]As shown in (12) after this end of read operation, and among drawing 8, the array controller 5 notifies file lead termination status to the central controlling part 2. To this status, the central controlling part 2 publishes the transfer request of the file B to the 2nd add-in board 9, as shown in (13) among drawing 8. In response to this transfer request, the 2nd add-in board 9 publishes file B transfer start status to the central controlling part 2, as shown in (14) among drawing 8. Transmission of a file is started by this as shown in (15) among drawing 8. An end of a file transfer will publish file transfer termination status to the central controlling part 2, as the 2nd add-in board 9 is shown in (16) among drawing 8 (16).

[0053]The above is explained with reference to drawing 7 like this embodiment about the operation in the case of having two I/F between the central controlling part 2 and the array controller 5, although it was a case of the comparative example to which the array controller 5 and the central controlling part 2 carry out data transfer in one I/F next.

[0054]The difference of timing with the case of this embodiment shown in the case of the comparative example shown in drawing 8 and drawing 7 is from the operation of (8) and (9) shown in a figure. That is, since I/F which connects the array controller 5 with the central controlling part 2 was one when drawing 8 mentioned above shows, "a notice of a lead demand command of the file B" shown in (9) among "the data transfer of the file A" shown in (8) among drawing 8 and drawing 8 was not able to be performed simultaneously. In the case of this embodiment, almost simultaneously with that (refer to (7) in drawing 7) which the data transfer of the file A by the side of the 2nd add-in board 9 starts via the DATA buffer 18 and 17 copies of DATA I/F of the array controller 5, to it. The request demand of the file B notified to the central controlling part 2 via the 2nd I/F part 4 from the 2nd add-in board 9 is notified to the array controller 5 via the 1st I/F part 3 (refer to (8) in drawing 7). Operation of (7) in this drawing 7 and (8) can be performed almost simultaneous. That is, it is because it has the 2nd I/F part 4 in data transfer, and it has the 1st I/F part 3 in the notice of a command, so it becomes possible to perform these two operations simultaneously.

[0055]In this embodiment, the 2nd I/F part 4 by the side of the host computer B and the DATA I/F part 17 by the side of the array controller 5 are prepared as only for DATA transmission, The 1st I/F part 3 by the side of the host computer B and the command I/F part 14 by the side of the array controller 5 were prepared as only for [, such as control commands and status, ] control-commands pertinent information. For this reason, as shown in drawing 7, the central controlling part 2, Via the 2nd I/F part 4, even if it is during the data transfer of the file A between the DATA I/F parts 17 by the side of the array controller 5, The read request of the file B notified from the 2nd add-in board 9 can be immediately notified via the command I/F part 14 by the side of the array controller 5 via the 1st I/F part 3. And to the request, within the array controller 5, as shown in (10) among drawing 7, the read operation of the file B is started, and as shown in (12) among drawing 7, file B lead termination status is notified to the central

controlling part 2 after the end of a lead.

[0056]In the 1st add-in board 8 side, transmission of the file A is continued regardless of operation of said 2nd add-in board 9, and as shown in (11) after the end of transmission, and among drawing 7, file A transmission termination status is notified from the 1st add-in board 8 to the central controlling part 2. Therefore the operation after the transfer-request command over the file B shown in (13) among drawing 7 is shown hereafter, it is the same as that of the case of the data transfer processing to the file A fundamentally mentioned above.

[0057]Thus, it constituted from a video server of this embodiment so that data transfer between the external devices through the central controlling part 2 or the add-in boards 8 and 9 by the side of the array controller 5 and the host computer B might be performed in two I/F. That is, the 2nd I/F part 4 by the side of a data control command or the host computer B and the DATA I/F part 17 by the side of the array controller 5 are prepared as only for DATA transmission, The 1st I/F part 3 by the side of the host computer B and the command I/F part 14 by the side of the array controller 5 were prepared as only for [, such as control commands and status, ] control-commands pertinent information.

[0058]For this reason, even if it is during the data transfer to an external device, for example, another control commands from the central controlling part 2 can be received, and data transfer high-speed as a whole can be realized. That is, since I/F for control-commands transmission and I/F for data transfer have dissociated, Even if the occupied state from the data transfer start of the I/F parts for data transfer (the DATA I/F part 17 by the side of the → array controller 5, the 2nd I/F part 4 by the side of the host computer B, etc.) to an end arises, Processing of control-commands transmission in the I/F part for control-commands transmission (the command I/F part 14 by the side of the → array controller 5 and the 1st I/F part 3 by the side of the host computer B) etc. is not affected.

[0059]When data is conversely supplied from an external device and it writes the data in the hard disk drive 7, it is effective similarly. And if the high-speed data transfer between the central controlling part 2 and the array controller 5 of it becomes possible as mentioned above since this system is a video server, its response time to the supply demand of the video data etc. which were carried out from the terminal unit of the data supply point will improve. That is, in the case of the system of a video server, it is possible that the request from two or more terminal units is advanced by arbitrary timing, but. If a host computer cannot take out directions to an array controller to the 2nd request unless the data transfer according to the first request is completed, the data transfer according to the 2nd request itself will be overdue. Similarly, the response as the whole — 3rd henceforth will be overdue — will fall. Therefore, if command sending out from the central controlling part 2 to the array controller 5 side is not influenced by the transmission processing for the data supply to an external terminal unit etc. as mentioned above with reference to drawing 7, The response in the portion improves and the response to the data supply demand carried out from the external terminal unit also improves.

[0060]When it applies to the system of the video-on-demand (VOD) service which enabled it to see the movie etc. which are wished to have when a user wishes especially, Since it is what outputs data at any time according to the request from a user's terminal unit, that improvement in a response to a data supply demand can be aimed at leads to the progression in quality of data supply service.

[0061]As mentioned above, this invention is not limited to such an example at all, and can be carried out with the gestalt which becomes various in the range which does not deviate from the main point of this invention. For example, although the above-mentioned embodiment explained the case where it realized as a video server, If it is a system for which the function of transmission is needed outside in the data which command sending out in the array controller 5 from the host computer B (in this case, central controlling part 2) and the array controller 5 read from the hard disk drive 7, it will not be limited to a video server.

[0062]Although the above-mentioned embodiment explained the thing based on a SCSI standard as a premise, it is not limited to such a standard.

[Translation done.]



## 【特許請求の範囲】

【請求項1】 複数台の記憶装置を並列に制御可能であり、外部のホストコンピュータからの指示に基づき、前記記憶装置から読み出したデータを前記ホストコンピュータを含む外部装置へ転送可能なアレイコントローラにおいて、

前記ホストコンピュータとの間で制御コマンド関連情報の転送を行なうためのコマンド用インタフェース部と、当該コマンド用インタフェース部とは別個に設けられ、前記外部装置に対して前記記憶装置から読み出したデータの転送を行なうためのデータ用インタフェース部と、を備えることを特徴とするアレイコントローラ。

【請求項2】 前記ホストコンピュータからの指示に基づき、前記記憶装置へのデータの書き込みも可能であり、前記データ用インタフェース部は、前記外部装置からの書込用データの転送にも用いられるよう構成されていることを特徴とする請求項1に記載のアレイコントローラ。

【請求項3】 請求項1又は2に記載のアレイコントローラと、該アレイコントローラによって制御される複数の記憶装置と、前記アレイコントローラの前記コマンド用インタフェース部と接続するための第1のインタフェース部及び前記データ用インタフェース部と接続するための第2のインタフェース部を備えると共に、前記記憶装置アレイのアレイコントローラに対し、前記外部装置であるデータ供給先の端末装置からのリクエストに応じたデータを読み出して前記端末装置へ転送するよう指示可能であるホストコンピュータと、を備えることを特徴とするデータサーバ。

【請求項4】 請求項2に記載のアレイコントローラと、該アレイコントローラによって制御される複数の記憶装置と、前記アレイコントローラの前記コマンド用インタフェース部と接続するための第1のインタフェース部及び前記データ用インタフェース部と接続するための第2のインタフェース部を備えると共に、前記記憶装置アレイのアレイコントローラに対し、前記外部装置であるデータ供給先の端末装置からのリクエストに応じたデータを読み出して前記端末装置へ転送するよう指示可能であると共に、前記外部装置から転送されてくる書込用データを前記記憶装置へ書き込むよう指示可能であるホストコンピュータと、を備えることを特徴とするデータサーバ。

【請求項5】 前記アレイコントローラのコマンド用インタフェース部と接続するためのホストコンピュータ側の前記第1のインタフェース部として、ホストコンピ

ュータが備えているプリンタ用のインタフェース部又はRS232Cインタフェース部を用いたことを特徴とする請求項3又は4に記載のデータサーバ。

## 【発明の詳細な説明】

## 【0001】

【発明の属する技術分野】本発明は、複数の記憶装置を並列に動作させることで、例えばコンピュータ機器で処理される映像や音声などの膨大なデータを高速かつリアルタイムに記録させ、また読み出しさせることのできるアレイコントローラ及びそのアレイコントローラを備えたデータサーバに関する。

## 【0002】

【従来の技術】従来より、例えばハードディスク装置などの記憶装置を複数台並列に動作させることで、メインコンピュータからのデータ読み出しあるいはデータ書き込みの要求に高速かつリアルタイムに対応することのできる記憶装置アレイが知られており、その制御部となるアレイコントローラも知られている。また、この記憶装置アレイを利用するデータサーバも考えられている。

【0003】そして、従来のアレイコントローラには、ホストコンピュータとしてのパーソナルコンピュータ（以下、パソコンと略記する。）又はワークステーション用のインタフェース部で、PCI（Peripheral Component Interconnect）と呼ばれるインタフェース部を備えた拡張ボードの形態のものがある。この拡張ボード形態のアレイコントローラは、パソコン又はワークステーション側の拡張スロットに挿入され、アレイコントローラ自身は制御対象の記憶装置群と接続することで、パソコン等からの指示に基づき、記憶装置群に対する制御を行うものである。

## 【0004】

【発明が解決しようとする課題】ところで、ホスト装置としてのパソコンやワークステーション等とアレイコントローラとの間のインタフェース部においては、記憶装置群に記憶された読み出し用のデータや、記憶装置群に記録するための書込用データを高速に転送する必要がある。また、読み出したデータや書込用データ以外にも、データの読み出し処理や書き込み処理を行わせるための種々の制御コマンドや、制御コマンドに対する結果として処理対象から返答されるステータスなども転送する必要がある。

【0005】しかしながら、記憶装置群に記憶された読み出し用のデータは、例えばビデオデータに代表されるように一般にデータサイズが大きく、そのデータ転送には所定の時間が必要となる。したがって、そのデータ転送中にはホストコンピュータが新たなデータ読み出しの指示をアレイコントローラに出すことができない。つまり、データ伝送が終了するまで待つこととなりレスポンスは低下する。そしてこのレスポンスは、ホストコンピュータがデータ読み出し指示を出す頻度が高くなるにつ

れて低下していく。

【0006】そのため、データサーバに適用した場合、多数の端末からのリクエストを受け付け、そのリクエストに応じたデータを極力リアルタイムに供給しようとする場合には、上記ホストコンピュータとアレイコントローラ間のレスポンスの低下が、端末からのリクエストに対するデータ供給のレスポンスの低下につながっていくのである。例えば利用者が希望するときに希望する映画等を見ることができるようにしたビデオ・オン・デマンド（VOD）サービスが研究されており、このVODサ

ービスは、情報出力側となるユーザ側に設けられたセットトップボックス（STB）から所望のビデオ放映がリクエストされると、サービス提供者側に用意されたビデオサーバと呼ばれる情報供給装置から、そのリクエストに応じた映像・音声をユーザに送信する。このようなシステムでは、情報出力端末側に対して出力される供給用情報はビデオ情報であり、情報出力端末からのリクエストに応じて随時出力されるものである。そのため、ある一つのリクエストに対するデータ転送処理が終了しないと次の処理のための制御コマンドも発行できないとレスポンスが低下してしまい、このレスポンスの低下は情報供給サービスの質の低下につながってしまう。

【0007】本発明は、上述した問題点を解決するためになされたものであり、ホストコンピュータとの間での高速なデータ転送が可能なアレイコントローラを提供することを第1の目的とする。また、ビデオサーバなどに適用した場合には応答時間の向上の点で有効なデータサーバの提供を第2の目的とする。

【0008】

【課題を解決するための手段及び発明の効果】前記第1の目的を達成するためになされた請求項1に記載の発明は、複数台の記憶装置を並列に制御可能であり、外部のホストコンピュータからの指示に基づき、記憶装置から読み出したデータをホストコンピュータを含む外部装置へ転送可能なアレイコントローラにおいて、ホストコンピュータとの間で少なくとも制御コマンド関連情報の転送を行なうためのコマンド用インタフェース部と、コマンド用インタフェース部とは別個に設けられ、外部装置に対して記憶装置から読み出したデータの転送を行なうためのデータ用インタフェース部とを備えることを特徴とする。なお、制御コマンド関連情報には、制御コマンドそのものの他にも、その制御コマンドの結果として制御対象側より返答されるステータスのような情報が含まれる。

【0009】本発明のアレイコントローラは、外部のホストコンピュータからの指示に基づき、記憶装置から読み出したデータをホストコンピュータを含む外部装置へ転送することができるのであるが、ホストコンピュータとの間での制御コマンド関連情報、例えば制御コマンド

ェース部を介して行ない、記憶装置から読み出したデータを外部装置へ転送するのはデータ用インタフェース部を介して行なう。

【0010】このように、制御コマンドやステータスなどの制御コマンド関連情報の転送用のインタフェースと、記憶装置からデータを読み出す場合のデータ転送用のインタフェースとをそれぞれ個別に準備することによって、例えば外部装置へのデータ転送中であってもホストコンピュータからの別の制御コマンドを受けることができ、全体として高速なデータ転送が実現できるアレイコントローラを得ることができる。つまり、制御コマンド転送用のインタフェースとデータ転送用のインタフェースとが分離しているので、データ転送開始から終了までのデータ転送用のデータ用インタフェース部の占有状態が生じて、コマンド用インタフェース部は影響を受けず、ホストコンピュータとの間で制御コマンド制御部は所定のプログラムに基づく制御処理を中断することなく継続して実行することができる。

【0011】なお、記憶装置からのデータ読み出し・転送だけでなく、ホストコンピュータからの指示に基づき記憶装置へのデータの書き込みも可能に構成すると共に、データ用インタフェース部を、外部装置からの書込用データの転送にも用いるよう構成することも考えられる。このようにすれば、記憶装置から読み出して外部装置へ転送するデータ自体を更新・追加したりすることができ、その書き込み時においても高速なデータ転送が実現できる。

【0012】一方、第2の目的を達成するためになされた請求項3に示す発明は、上述したアレイコントローラと、アレイコントローラによって制御される複数の記憶装置と、アレイコントローラのコマンド用インタフェース部と接続するための第1のインタフェース部及びデータ用インタフェース部と接続するための第2のインタフェース部を備えると共に、記憶装置アレイのアレイコントローラに対し、外部装置であるデータ供給先の端末装置からのリクエストに応じたデータを読み出して端末装置へ転送するよう指示可能であるホストコンピュータとを備えることを特徴とするデータサーバである。

【0013】本データサーバは、いわゆるビデオサーバなどに適用することができるが、その場合には、データ供給先の端末装置からされたビデオデータの供給要求に対しての応答時間の向上につながる。この場合には、複数の端末装置からのリクエストが任意のタイミングで出されることが考えられるが、最初のリクエストに応じたデータ転送が終了しないと、2番目のリクエストに対してホストコンピュータがアレイコントローラに指示を出せないのでは、その2番目のリクエストに応じたデータ転送自体も遅れてしまう。同様に、3番目以降も遅れてしまうなど、全体としてのレスポンスが低下してしまう。したがって、上述したように、アレイコントローラ

とホストコンピュータ間でのやり取りがデータ転送には影響されないようにすることで、その部分でのレスポンスが向上し、外部の端末装置からされたデータ供給要求に対してのレスポンスも向上するのである。

【0014】特に、ビデオサーバに適用した場合には、利用者が希望するときに希望する映画等を見ることができるようにしたビデオ・オン・デマンド（VOD）サービスのシステムに利用されることが考えられる。このようなシステムでは、利用者の端末装置からのリクエストに応じて随時データを出力するものであるため、データ供給要求に対してのレスポンスの向上を図ることができることは、データ供給サービスの質の向上につながるものである。

【0015】なお、請求項4に示すように、記憶装置へのデータ書き込み制御もできるアレイコントローラを用いる場合には、ホストコンピュータを、外部装置から転送されてくる書込用データを記憶装置へ書き込むよう指示可能に構成すればよい。この場合も、同様に、アレイコントローラとホストコンピュータ間でのやり取りがデータ転送には影響されないためその部分でのレスポンスが向上し、記憶装置への書込用データの書き込み動作などがあっても、やはり外部の端末装置からされたデータ供給要求に対してのレスポンスが向上する。

【0016】なお、このようなデータサーバとして実現する場合、ホストコンピュータ側におけるデータ転送用の第2のインタフェース部としては従来技術としても説明したPCI用のI/Fを用いることができる。これをパソコン等の拡張スロットに導入することで接続し、データ転送用とする。また、制御コマンド関連情報の転送用のコマンド用インタフェース部に対応するホストコンピュータ側の第1のインタフェース部としては、例えばホストコンピュータが備えているプリンタ用のインタフェース部又はRS232Cインタフェースを用いることが考えられる。もちろん、専用のインタフェース部を設けてもよいが、プリンタ用のインタフェース部又はRS232Cインタフェースであれば元々備えているコンピュータが多いので、それを利用することができる。

#### 【0017】

【発明の実施の形態】以下、本発明の実施の形態について図面を参照して説明する。図1は、ビデオサーバに適用した場合の全体の構成を示すブロック図である。アレイコントローラ5には、複数のSCSIバス6が接続されている。なお、複数のSCSIバス6を区別する場合には、SCSIバス(0)～SCSIバス(m)として記す。また、その個々のSCSIバス6には複数のハードディスク装置(HDD)が接続されている。なお、複数のハードディスク装置7を区別する場合にはID

(0)～ID(n)を付して記すこととする。これら、アレイコントローラ5、複数のSCSIバス6及びそれに接続される複数のハードディスク装置7によってディ

スクアレイAが構成されている。

【0018】一方、このディスクアレイAを制御するホストコンピュータBは、記憶部1、中央制御部2、第1のインタフェース（以下「I/F」と略記する）部3、第2のI/F部4、第1の拡張ボード8及び第2の拡張ボード9などによって構成されている。

【0019】第1のI/F部3は「制御コマンド関連情報」としての制御コマンドやステータスの転送を行うためのインタフェースであり、第2のI/F部4は主にデータ転送を行うためのインタフェースである。上述したアレイコントローラ5は、これら2つのI/F部3、4を介して中央制御部2に接続されている。

【0020】また、第2のI/F部4には、第1の拡張ボード8及び第2の拡張ボード9が接続されている。これら2つの拡張ボード8、9は、ディスクアレイAのハードディスク装置7に書き込むデータを外部から受信し、その受信データをアレイコントローラ5に転送すると共に、ハードディスク装置7から読み出されアレイコントローラ5内の後述するバッファに記録されたデータを、図示しない通信回線を介して接続された端末装置（図示せず）などに対して送信するためのLANボードである。

【0021】ホストコンピュータBの中央制御部2は、ビデオサーバとしての主な動作の制御主体となるが、大きく分けて次の2種類がある。その一つは、第1の拡張ボード8又は第2の拡張ボード9に収容された通信回線（図示せず）を介して接続された端末装置（図示せず）からリクエストが出された場合に、その端末装置に対してリクエストに応じたデータを送信するという動作であり、もう一つは、新規に音声及び映像データを外部から取り込み、それをディスクアレイAのハードディスク装置7に記憶させるという動作である。

【0022】外部の端末装置へのデータ配信は以下のように行われる。ビデオサーバの制御主体としての中央制御部2は、第1の拡張ボード8及び第2の拡張ボード9を介して接続された例えば端末装置としてのSTB（Set Top Box）等から送られた映像や音声データの供給リクエストに対して、ハードディスク装置7にファイルとして記憶されたリクエストに応じたデータを、アレイコントローラ5を介して読み出し、第1の拡張ボード8及び第2の拡張ボード9を介して端末装置側へ供給する。

【0023】一方、ハードディスク装置7へデータ記憶動作は以下のように行われる。新規な音声及び映像データが第1の拡張ボード8や第2の拡張ボード9を介して外部から取り込まれたり、あるいは第2のI/F部4に接続された図示しないI/F部等を介して取り込まれ、アレイコントローラ5へ転送される。アレイコントローラ5は、その転送されたデータをファイル等の形式でハードディスク装置7へ記録する。

【0024】次に、図2を参照してアレイコントローラ

5について説明する。アレイコントローラ5は、アレイコントローラ5全体の制御を司るCPU11と、プログラムや各種パラメータ等の記憶されたROM12と、CPU11のワークエリアとして、またデータの記憶エリアとして使用されるRAM13と、「コマンド用I/F部」に相当するコマンドI/F部14と、割込制御部15と、複数のコマンドバッファ16と、「データ用I/F部」に相当するDATA I/F部17と、複数のDATAバッファ18と、リード/ライト・パリティ生成及びバス切換部19と、2種類のチャックメモリ、すなわちチャックメモリ(1)20及びチャックメモリ(2)21と、SCSI制御部22とを備えている。なお、チャックメモリ(1)20及びチャックメモリ(2)21とSCSI制御部22は、SCSIバス6毎に設けられており、本実施形態ではSCSIバス6がID(0)～(m)まで存在するので、同様にSCSI制御部22もID(0)～(m)まで存在し、それぞれにチャックメモリ(1)20及びチャックメモリ(2)21が接続されることとなる。

【0025】前記コマンドI/F部14は、ホストコンピュータB側の第1のI/F部3(図1参照)を介して中央制御部2に接続されると共に、アレイコントローラ5内のコマンドバッファ16と割込制御部15とに接続され、制御コマンド関連情報の転送に使用される。この「制御コマンド関連情報」には、制御コマンドそのものの他にも、その制御コマンドの結果として制御対象側より返答されるステータスのような情報が含まれる。

【0026】コマンドバッファ16は、1つ又は複数のコマンドを同時に処理できるよう複数エリア分用意されている。なお、複数のコマンドバッファ16を区別する場合にはコマンドバッファ(1)～コマンドバッファ(h)というように1～hの番号を付して記すこととする。

【0027】また、割込制御部15は、コマンドバッファ16にコマンドをセットした後、中央制御部2がアレイコントローラ5に対して割込を発生させる場合にはCPU11に対して、またアレイコントローラ5が中央制御部2に対して割込を発生させる場合には第1のI/F部3を介して、それぞれ中央制御部2に対して割込を発生する機能を備えている。

【0028】なお、ホストコンピュータBとしていわゆるパーソナルコンピュータを適用する場合、例えば第1のI/F部3としては、プリンタポートやRS232Cポート等汎用のポートを適用してもよいし、もちろん専用のI/Fを適用することも可能である。

【0029】一方、DATA I/F部17は、ホストコンピュータB側の第2のI/F部4(図1参照)を介して中央制御部2に接続され、また第1の拡張ボード8及び第2の拡張ボード9等に接続されると共に、アレイコントローラ5内のDATAバッファ18に接続され、主

にデータの転送時に使用される。

【0030】DATAバッファ18は、複数の拡張ボード8、9からの要求に応えられるよう複数エリア分用意されている。なお、複数のDATAバッファ18を区別する場合にはDATAバッファ(1)～コマンドバッファ(k)というように1～kの番号を付して記すこととする。

【0031】なお、このDATA I/F部17としては従来技術としても説明したPCI(Peripheral Component Interconnect)の技術を用いることができる。つまり、アレイコントローラ5自体が拡張ボード形態であり、この拡張ボード形態のアレイコントローラ5をホストコンピュータB側の拡張スロットに挿入して接続することとなる。この際のホストコンピュータB側の第2のI/F部4も、もちろんPCIに対応したI/Fである。

【0032】次に、リード/ライト・パリティ生成及びバス切換部19の動作について説明する。リード/ライト・パリティ生成及びバス切換部19は、図3～図6に示すように、内部バス側入出力制御部30、パリティ生成及びチェック部31及びSCSIバス側入出力制御部32とを備えている。

【0033】本実施形態では、内部バスを32ビット幅と想定し、8ビット毎にまとめた単位のバスとして扱い、それぞれのバスを識別のための0～3の番号を付してAバス0、Aバス1、Aバス2、Aバス3と表す。また、アレイコントローラ5に接続されるSCSIバスを、5系統の8ビットのSCSIバスとし、その中の1系統は、他の4系統のパリティ用のバスであるものとして説明する。

【0034】最初に、ハードディスク装置7に対するライト時の動作について説明する。図3に示すように、このときの内部バス側入出力制御部30は、Aバス0～3に対しては入力となり、Bバス0～3に対しては出力となる。詳しくは、Bバス0にはAバス0、Bバス1にはAバス1、Bバス2にはAバス2、Bバス3にはAバス3の内容がそのまま出力されることとなる。また、パリティ生成及びチェック部31には、4本のBバス0～3の値が入力され、その4つの値のパリティ計算値がBバスPに出力される。SCSIバス側入出力制御部32では、BバスPとBバス0～3の値が入力され、Cバス0～4に出力される。

【0035】次に、ハードディスク装置7からの通常のリード時の動作について図4を参照して説明する。SCSIバス側入出力制御部32は、ハードディスク装置7からのリードの動作であるため、Cバス0～4側が入力となり、そのバスの値が、BバスP及びBバス0～3に出力される。パリティ生成及びチェック部31には、BバスP及びBバス0～3の値が入力されるため、Bバス0～3の値の排他的論理和をとることによりパリティ値



を計算し、BバスPの値と比較した結果をパリティチェック結果として出力する。また、内部バス側入出力制御部30にはBバス0～3が入力され、Aバス0～3に出力する。

【0036】次に、識別番号3のバスに接続されたハードディスク装置7が故障し、識別番号0のバスに接続されたハードディスク装置7と他の3系統（識別番号1～3）のバスに接続されたハードディスク装置7とから、識別番号4のバスに接続されたハードディスク装置7のデータが復元される場合の動作について、図5を参照して説明する。

【0037】この場合、SCSIバス側入出力制御部32には、Cバス0～3の値が入力され、Cバス4の値は不定となっている。また、Bバス0～3側に対しては、BバスP及びBバス0～2に対してCバス0～3の値がそれぞれ出力される。そして、パリティ生成及びチェック部31にはBバスP及びBバス0～2の値が入力され、それらの値の排他的論理和をとることによってCバス4に出力されるべき値が復元され、Bバス3に出力される。内部バス側入出力制御部30には図4の場合と同様にBバス0～3が入力され、そのままAバス0～3へ出力される。

【0038】次に、識別番号4のバスに接続されたハードディスク装置7が故障し、故障したハードディスク装置7が正常なハードディスク装置7に交換された後、識別番号0～3のバスに接続されたハードディスク装置7からのデータによって、識別番号4のバスに接続されていたハードディスク装置7のデータが再構築（リビルド）される場合の動作について図6を参照して説明する。

【0039】SCSIバス側入出力制御部32には、Cバス0～3の値が入力され、Bバス0～3側に対しては、BバスP及びBバス0～2に対してCバス0～3の値がそれぞれ出力される。パリティ生成及びチェック部31にはBバスP及びBバス0～2の値が入力され、それらの値の排他的論理和をとることによってCバス4に接続されていたハードディスク装置7に記憶されていたデータが復元され、Bバス3に出力される。

【0040】SCSIバス側入出力制御部32では、Bバス3のみが他のバスと異なり入力となり、Cバス4に対してデータが出力される。これによって再生されたデータをハードディスク装置7に書き込み、故障したハードディスク装置7と全く同じ内容のデータを再構築することが可能となる。またこの時、内部バス側入出力制御部30のBバス0～3側を入力とし、Aバス0～3側からその値を出力することにより、ハードディスク装置7に記憶されたデータを再構築しながら、同時にリードすることも可能となる。

【0041】上記の動作により、リード/ライト・パリティ生成及びバス切換部19を介して、DATAバッ

ファ18と第1のチャックメモリ20又は第2のチャックメモリ21（図1参照）との間で、リードデータ及びライトデータの転送が行われる。

【0042】次に、チャックメモリ20、21からハードディスク装置7へデータを転送動作と、その逆にハードディスク装置7からチャックメモリ20、21へデータを転送する動作について説明する。上述したように、各SCSIバス6にはハードディスク装置7が1又は複数個接続されている。SCSI制御部22は、SCSIバス6の制御を行うと共に、書込時には、チャックメモリ（1）20及びチャックメモリ（2）21に書き込まれたデータをハードディスク装置7に転送し、読み出し時には、ハードディスク装置7から読み出されたデータを、チャックメモリ（1）20又はチャックメモリ（2）21に転送する。チャックメモリが2つある理由は、ハードディスク装置7に対するアクセスを高速化するためである。

【0043】以下、書込時及び読み出し時のデータの流れをまとめて説明する。書込時には、DATA I/F部17を介して、複数個（1～k）用意されたDATAバッファ18の内いずれかのDATAバッファ18に書込データが転送される。DATAバッファ18からは、リード/ライト・パリティ生成部及びバス切換部19を介してチャックメモリ（1）20へ書込データが転送される。転送が終了すると、チャックメモリ（1）20からSCSI制御部22を介してSCSIバス6に接続されたハードディスク装置7にデータが転送され、記憶される。このデータ転送と同時に、もう一方のチャックメモリ（2）21に、次にハードディスク装置7転送すべきデータをDATAバッファ18から予め転送しておく。これによって、1つのチャックメモリからハードディスク装置7へのデータ転送終了後、無駄な時間をあけることなく、即座にもう1つのチャックメモリのデータをハードディスク装置7へ転送することができ、書込動作の高速化が可能となる。

【0044】また、読み出し時には、その逆に、ハードディスク装置7→SCSIバス6→SCSI制御部22→チャックメモリ（1）20又はチャックメモリ（2）21の順でデータが転送される。読み出し時も書込時と同様に、チャックメモリ20、21からDATAバッファ18へのデータ転送中に、空いたチャックメモリ20、21に対してハードディスク装置7から次の読み出しデータを転送する。それによって、読み出し動作を高速化することができる。

【0045】次に、2つの拡張ボード8、9からファイルの転送要求があった場合の動作について、図7及び図8のシーケンス図を参照して説明する。図7は、本実施形態のアレイコントローラ5の場合であって、図2に示すコマンドI/F14及びDATA I/F17を介し、それぞれホストコンピュータB側の第1のI/F部3及

び第2のI/F部4との間でデータ転送を行なう場合の例である。一方、図8は、比較のために、アレイコントローラ5がI/Fを一つしか持たず、したがって中央制御部2と間で一つのI/Fにてデータ転送を行なう場合の例である。

【0046】まず、図8を参照して、一つのI/Fにてアレイコントローラ5と中央制御部2とがデータ転送を行なう比較例の場合について説明する。図8中(1)に示すように、最初に、第1の拡張ボード8からファイルAの転送が中央制御部2にリクエストされる。図8中(2)に示すように、中央制御部2はアレイコントローラ5に対してファイルAのリード要求を発行する。

【0047】すると、図8中(3)に示すように、アレイコントローラ5内では、その要求に対してハードディスク装置7からファイルAのデータ読み込み、前記動作手順に従って、SCSI制御部22、チャックメモリ20、21及びリード/ライト・パリティ生成及びバス切換部19を介してDATAバッファ18にデータを転送する。

【0048】図8中(4)に示すように、アレイコントローラ5は、ファイルリード終了後にファイルAのリード終了を中央制御部2に対して通知する。そして、図8中(5)に示すように、中央制御部2では、ファイルAリード終了を受けた後、第1の拡張ボード8に対してファイルAの転送要求を発行する。そして、図8中(6)に示すように、第1の拡張ボード8では、その要求に対してファイルAの転送開始を中央制御部2に通知する。

【0049】この直後、図8中(7)に示すように、第2の拡張ボード9からファイルBに対するリクエストが発生し、中央制御部2に通知される。しかし、このときは第1の拡張ボード8によって中央制御部2に対してファイルAの転送開始ステータスが通知された後であるため、図8中(8)に示すように、第1の拡張ボード8はファイルAの転送を開始する。

【0050】図8中(8)で開始する第1の拡張ボード8によるデータ転送中は、バスが、アレイコントローラ5内のDATAバッファ18から拡張ボード8へのデータ転送に使用されている。このため、中央制御部2は、第2の拡張ボード9からのファイルBのリクエストをすぐ発行することができない。そのため、図8中(9)に示すように、中央制御部2は、アレイコントローラ5から第1の拡張ボード8へのファイルAの転送が終了した後に初めてアレイコントローラ5に対してファイルBのリード要求を通知することとなる。

【0051】そして、図8中(10)に示すように、中央制御部2では、ファイルAの転送終了ステータスを第1の拡張ボード8から受け取る。また、図8中(11)に示すように、それとほぼ同時に、アレイコントローラ5ではファイルBのリード要求に対してハードディスク装置7からファイルリード動作を開始し、DATAバッ

ファ18に、ハードディスク装置7から読み込んだファイルBのデータを転送する。

【0052】このリード動作終了後、図8中(12)に示すように、アレイコントローラ5は中央制御部2に対してファイルリード終了ステータスを通知する。このステータスに対して、中央制御部2は、図8中(13)に示すように、第2の拡張ボード9に対してファイルBの転送要求を発行する。この転送要求を受けて、第2の拡張ボード9は、図8中(14)に示すように、ファイルB転送開始ステータスを中央制御部2に対して発行する。これによって、図8中(15)に示すようにファイルの転送が開始される。ファイル転送が終了すると、第2の拡張ボード9は、図8中(16)に示すように、ファイル転送終了ステータスを中央制御部2に対して発行する(16)。

【0053】以上は、一つのI/Fにてアレイコントローラ5と中央制御部2とがデータ転送を行なう比較例の場合であったが、次に、本実施形態のように、中央制御部2とアレイコントローラ5との間に2系統のI/Fを持つ場合の動作について、図7を参照して説明する。

【0054】図8に示す比較例の場合と図7に示す本実施形態の場合とのタイミングの相違は、図中に示す

(8)、(9)の動作からである。つまり、上述した図8で示す場合には、中央制御部2とアレイコントローラ5を接続するI/Fが1つであるため、図8中(8)に示す「ファイルAのデータ転送」と図8中(9)に示す「ファイルBのリード要求コマンド通知」を同時に行うことができなかった。それに対して、本実施形態の場合には、アレイコントローラ5のDATAバッファ18及びDATA I/F 17部を介して第2の拡張ボード9側へのファイルAのデータ転送が開始する(図7中(7)参照)のとほぼ同時に、第2の拡張ボード9から第2のI/F部4を介して中央制御部2に通知されたファイルBのリクエスト要求が、第1のI/F部3を介してアレイコントローラ5に対して通知される(図7中(8)参照)。この図7中の(7)、(8)の動作はほぼ同時に行うことができる。つまり、データ転送用に第2のI/F部4を、コマンド通知用に第1のI/F部3を持っているので、これら2つの動作を同時に行うことが可能となるからである。

【0055】本実施形態では、ホストコンピュータB側の第2のI/F部4及びアレイコントローラ5側のDATA I/F部17をDATA転送専用として準備し、ホストコンピュータB側の第1のI/F部3及びアレイコントローラ5側のコマンドI/F部14を制御コマンドやステータスなどの制御コマンド関連情報専用として準備した。このため、図7に示すように、中央制御部2は、第2のI/F部4を介してアレイコントローラ5側のDATA I/F部17との間でのファイルAのデータ転送中であっても、第2の拡張ボード9から通知された

ファイルBのリードリクエストを、第1のI/F部3を介して即座にアレイコントローラ5側のコマンドI/F部14を介して通知することができる。そして、そのリクエストに対して、アレイコントローラ5内では、図7中(10)に示すように、ファイルBのリード動作が開始され、図7中(12)に示すように、リード終了後にファイルBリード終了ステータスが中央制御部2に対して通知される。

【0056】また、第1の拡張ボード8側では、前記第2の拡張ボード9の動作とは無関係にファイルAの転送が続けられ、転送終了後、図7中(11)に示すように、第1の拡張ボード8から中央制御部2に対してファイルA転送終了ステータスが通知される。以下、図7中(13)に示すファイルBに対する転送要求コマンド以降の動作を示すについては、基本的には上述したファイルAに対するデータ転送処理の場合と同様である。

【0057】このように、本実施形態のビデオサーバでは、アレイコントローラ5とホストコンピュータB側の中央制御部2あるいは拡張ボード8、9を介した外部装置との間でのデータ転送を2系統のI/Fにて行なうように構成した。つまり、データ制御コマンドやホストコンピュータB側の第2のI/F部4及びアレイコントローラ5側のDATA I/F部17をDATA転送専用として準備し、ホストコンピュータB側の第1のI/F部3及びアレイコントローラ5側のコマンドI/F部14を制御コマンドやステータスなどの制御コマンド関連情報専用として準備したのである。

【0058】このため、例えば外部装置へのデータ転送中であっても中央制御部2からの別の制御コマンドを受けことができ、全体として高速なデータ転送が実現できる。つまり、制御コマンド転送用のI/Fとデータ転送用のI/Fとが分離しているので、データ転送用I/F部(→アレイコントローラ5側のDATA I/F部17やホストコンピュータB側の第2のI/F部4など)のデータ転送開始から終了までの占有状態が生じても、制御コマンド転送用のI/F部(→アレイコントローラ5側のコマンドI/F部14やホストコンピュータB側の第1のI/F部3)での制御コマンド転送等の処理には影響を与えないのである。

【0059】なお、逆に外部装置からデータを供給され、そのデータをハードディスク装置7に書き込むような場合にも同様に有効である。そして、本システムはビデオサーバであるため、上述したように中央制御部2とアレイコントローラ5との間での高速なデータ転送が可能となれば、データ供給先の端末装置からされたビデオデータなどの供給要求に対しての応答時間が向上する。つまり、ビデオサーバのシステムの場合には、複数の端末装置からのリクエストが任意のタイミングで出されることが考えられるが、最初のリクエストに応じたデータ転送が終了しないと、2番目のリクエストに対してホス

トコンピュータがアレイコントローラに指示を出せないのでは、その2番目のリクエストに応じたデータ転送自体も遅れてしまう。同様に、3番目以降も遅れてしまうなど、全体としてのレスポンスが低下してしまう。したがって、図7を参照して上述したように、中央制御部2からアレイコントローラ5側へのコマンド送出が外部の端末装置などへのデータ供給のための転送処理に影響されないのであれば、その部分でのレスポンスが向上し、外部の端末装置からされたデータ供給要求に対してのレスポンスも向上するのである。

【0060】特に、利用者が希望するときに希望する映画等を見ることができるようにしたビデオ・オン・デマンド(VOD)サービスのシステムに適用した場合、利用者の端末装置からのリクエストに応じて随時データを出力するものであるため、データ供給要求に対してのレスポンスの向上を図ることができることは、データ供給サービスの質の向上につながる。

【0061】以上、本発明はこのような実施例に何等限定されるものではなく、本発明の主旨を逸脱しない範囲において種々なる形態で実施し得る。例えば、上記実施形態では、ビデオサーバとして実現した場合を説明したが、ホストコンピュータB(この場合は中央制御部2)からアレイコントローラ5へのコマンド送出及びアレイコントローラ5がハードディスク装置7から読み出したデータを外部へ転送という機能が必要とされるシステムであれば、ビデオサーバには限定されない。

【0062】また、上述の実施形態ではSCSI規格に準拠したものを前提として説明したが、そのような規格に限定されるものではない。

#### 【図面の簡単な説明】

【図1】 実施形態のビデオサーバ全体の構成を示すブロック図である。

【図2】 実施形態のアレイコントローラの内部構成を示すブロック図である。

【図3】 アレイコントローラ内のリード/ライト・パリティ生成及びバス切換部が実行するライト動作を示すブロック図である。

【図4】 アレイコントローラ内のリード/ライト・パリティ生成及びバス切換部が実行するリード動作を示すブロック図である。

【図5】 アレイコントローラ内のリード/ライト・パリティ生成及びバス切換部が実行するハードディスク装置状態でのリード動作を示すブロック図である。

【図6】 アレイコントローラ内のリード/ライト・パリティ生成及びバス切換部が実行するデータ再構築動作を示すブロック図である。

【図7】 実施形態のアレイコントローラと中央制御部あるいは拡張ボードを介した外部装置とのデータ転送を示すシーケンス図である。

【図8】 アレイコントローラと中央制御部あるいは拡張

張ボードを介した外部装置とのデータ転送を一つのI/Fにて実行する比較例の場合のデータ転送を示すシーケンス図である。

【符号の説明】

1…記憶部  
2…中央制御部  
3…第1のI/F部  
4…第2のI/F部  
5…アレイドコントローラ  
6…SCSIバス  
7…ハードディスク装置  
8…第1の拡張ボード  
9…第2の拡張ボード  
10…コマンドI/F部

\* 15…割込制御部

バッファ

17…DATA I/F部

バッファ

19…リード/ライト・パリティ生成及びバス切換部

20…チャンクメモリ(1)

メモリ(2)

22…SCSI制御部

側入出力制御部

31…パリティ生成及びチェック部

バス側入出力制御部

A…ディスクアレイ

\* ピュータ

16…コマンド

18…DATA

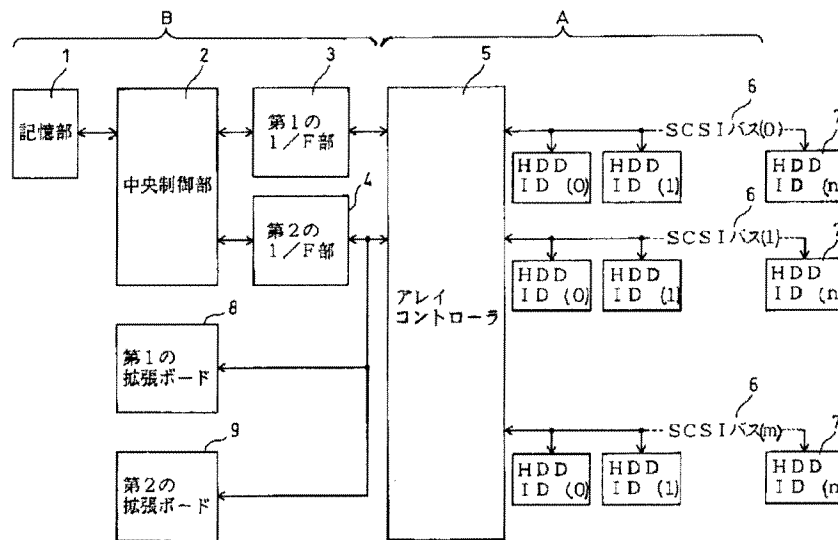
21…チャンク

30…内部バス

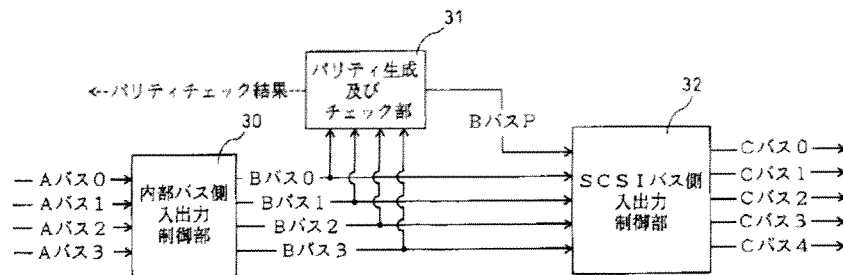
32…SCSI

B…ホストコン

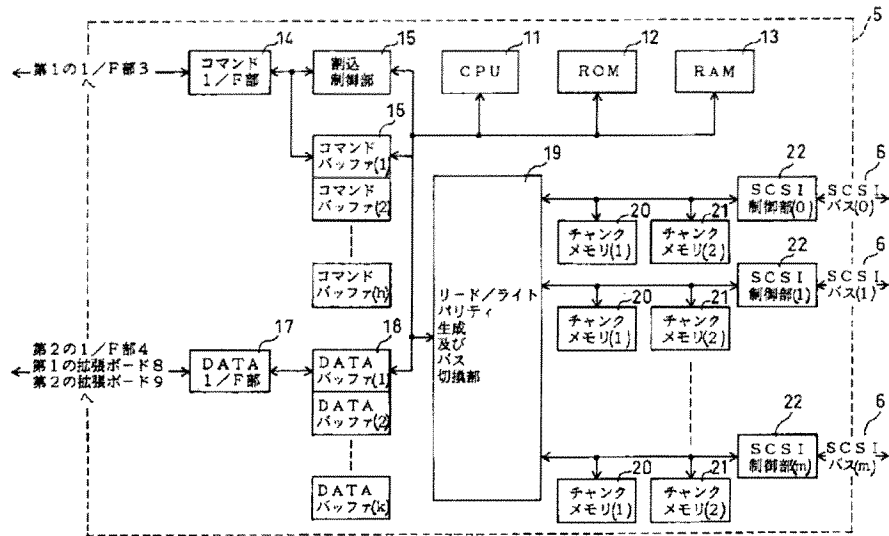
【図1】



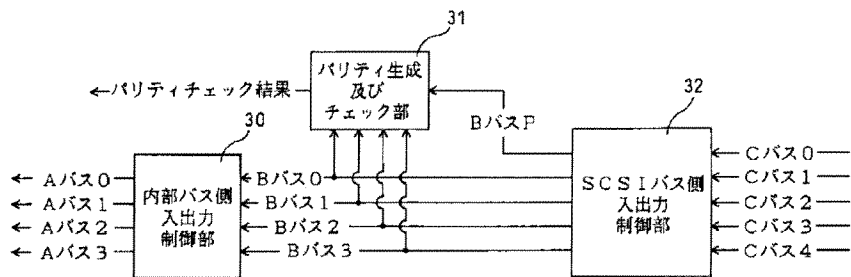
【図3】



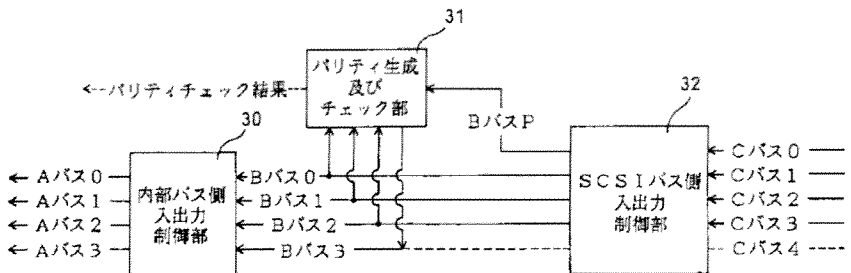
【図2】



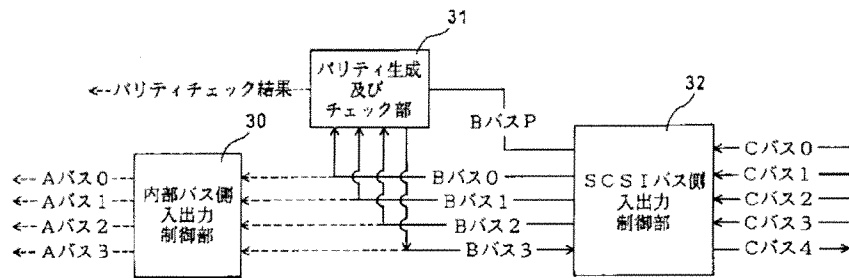
【図4】



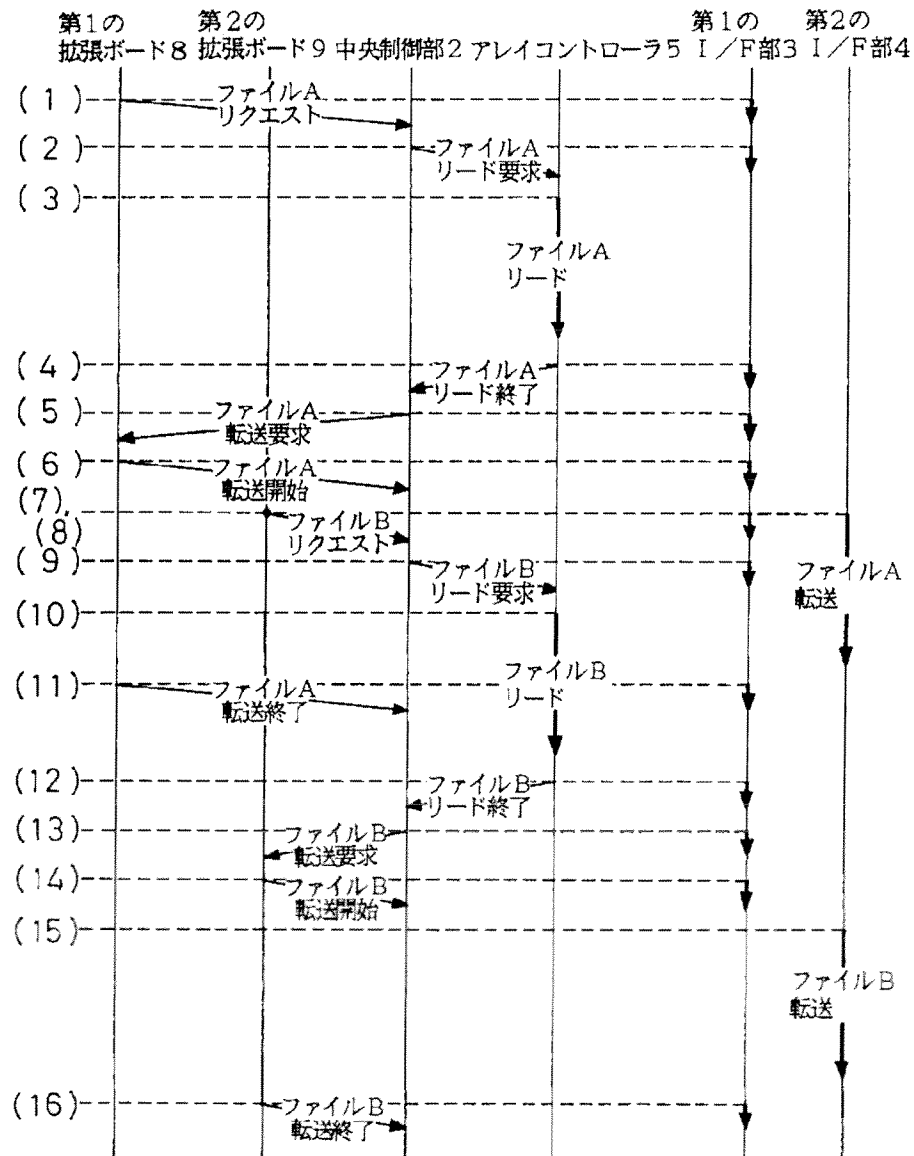
【図5】



【図6】



【図7】



【図8】

